CLAIMS

1. An image processing apparatus for estimating a motion of a predetermined feature point of a 3D object from a motion picture of the 3D object taken by a monocular camera, comprising:

observation vector extracting means for extracting projected coordinates of the predetermined feature point onto an image plane, from each of frames of the motion picture;

3D model initializing means for making the observation vector extracting means extract from an initial frame of the motion picture, initial projected coordinates in a model coordinate arithmetic expression for calculation of model coordinates of the predetermined feature point on the basis of a first parameter, a second parameter, and the initial projected coordinates; and

motion estimating means for calculating estimates of state variables including a third parameter in a motion arithmetic expression for calculation of coordinates of the predetermined feature point at a time of photography when a processing target frame of the motion picture different from the initial frame was taken, from the model coordinates, the first parameter, and the second parameter, and for outputting an output value about the motion of the predetermined feature point on the basis of the second parameter included in the estimates of the state variables,

wherein the model coordinate arithmetic expression is based on back projection of the monocular camera, the first parameter is a parameter independent of a local motion of a portion including the predetermined feature point, and the second parameter is a parameter

dependent on the local motion of the portion including the predetermined feature point, and

wherein the motion estimating means:

calculates predicted values of the state variables at the time of photography when the processing target frame was taken, based on a state transition model;

applies the initial projected coordinates, and the first parameter and the second parameter included in the predicted values of the state variables, to the model coordinate arithmetic expression to calculate estimates of the model coordinates at the time of photography;

applies the third parameter in the predicted values of the state variables and the estimates of the model coordinates to the motion arithmetic expression to calculate estimates of coordinates of the predetermined feature point at the time of photography;

applies the estimates of the coordinates of the predetermined feature point to an observation function based on an observation model of the monocular camera to calculate estimates of an observation vector of the predetermined feature point;

makes the observation vector extracting means extract the projected coordinates of the predetermined feature point from the processing target frame, as the observation vector; and

filters the predicted values of the state variables by use of the extracted observation vector and the estimates of the observation vector to calculate estimates of the state variables at the time of photography.

2. The image processing apparatus according to Claim 1, wherein the first parameter is a static parameter to converge at a specific

value, and wherein the second parameter is a dynamic parameter to vary with the motion of the portion including the predetermined feature point.

3. The image processing apparatus according to Claim 2, wherein the static parameter is a depth from the image plane to the predetermined feature point.

4. The image processing apparatus according to Claim 2 or 3, wherein the dynamic parameter is a rotation parameter for specifying a rotation motion of the portion including the predetermined feature point.

5. The image processing apparatus according to Claim 4, wherein the rotation parameter is angles made by a vector from an origin to the predetermined feature point, relative to two coordinate axes in a coordinate system whose origin is at a center of the portion including the predetermined feature point.

6. The image processing apparatus according to Claim 1, wherein the first parameter is a rigid parameter, and wherein the second parameter is a non-rigid parameter.

7. The image processing apparatus according to Claim 6, wherein the rigid parameter is a depth from the image plane to the model coordinates.

8. The image processing apparatus according to Claim 6 or 7, wherein the non-rigid parameter is a change amount about a position change of the predetermined feature point due to the motion of the portion including the predetermined feature point.

9. The image processing apparatus according to any one of Claims 1 to 8, wherein the motion model is based on rotation and

translation motions of the 3D object, and wherein the third parameter is a translation parameter for specifying a translation amount of the 3D object and a rotation parameter for specifying a rotation amount of the 3D object.

10. The image processing apparatus according to any one of Claims 1 to 9, wherein the motion estimating means applies extended Kalman filtering as said filtering.

11. An image processing method of estimating a motion of a predetermined feature point of a 3D object from a motion picture of the 3D object taken by a monocular camera, comprising:

a 3D model initialization step of extracting initial projected coordinates in a model coordinate arithmetic expression for calculation of model coordinates of the predetermined feature point on the basis of a first parameter, a second parameter, and the initial projected coordinates, from an initial frame of the motion picture; and

a motion estimation step of calculating estimates of state variables including a third parameter in a motion arithmetic expression for calculation of coordinates of the predetermined feature point at a time of photography when a processing target frame of the motion picture different from the initial frame was taken, from the model coordinates, the first parameter, and the second parameter, and outputting an output value about the motion of the predetermined feature point on the basis of the second parameter included in the estimates of the state variables,

wherein the model coordinate arithmetic expression is based on back projection of the monocular camera, the first parameter is a

parameter independent of a local motion of a portion including the predetermined feature point, and the second parameter is a parameter dependent on the local motion of the portion including the predetermined feature point,

wherein the motion estimation step comprises:

calculating predicted values of the state variables at the time of photography when the processing target frame was taken, based on a state transition model;

applying the initial projected coordinates, and the first parameter and the second parameter included in the predicted values of the state variables, to the model coordinate arithmetic expression to calculate estimates of the model coordinates at the time of photography;

applying the third parameter in the predicted values of the state variables and the estimates of the model coordinates to the motion arithmetic expression to calculate estimates of coordinates of the predetermined feature point at the time of photography;

applying the estimates of the coordinates of the predetermined feature point to an observation function based on an observation model of the monocular camera to calculate estimates of an observation vector of the predetermined feature point;

extracting projected coordinates of the predetermined feature point from the processing target frame, as the observation vector; and

filtering the predicted values of the state variables by use of the extracted observation vector and the estimates of the observation vector to calculate estimates of the state variables at the time of photography.

12. The image processing method according to Claim 11,

wherein the first parameter is a static parameter to converge at a specific value, and wherein the second parameter is a dynamic parameter to vary with the motion of the portion including the predetermined feature point.

13. The image processing method according to Claim 12, wherein the static parameter is a depth from the image plane to the predetermined feature point.

14. The image processing method according to Claim 12 or 13, wherein the dynamic parameter is a rotation parameter for specifying a rotation motion of the portion including the predetermined feature point.

15. The image processing method according to Claim 14, wherein the rotation parameter is angles made by a vector from an origin to the predetermined feature point, relative to two coordinate axes in a coordinate system whose origin is at a center of the portion including the predetermined feature point.

16. The image processing method according to Claim 11, wherein the first parameter is a rigid parameter, and wherein the second parameter is a non-rigid parameter.

17. The image processing method according to Claim 16, wherein the rigid parameter is a depth from the image plane to the model coordinates.

18. The image processing method according to Claim 16 or 17, wherein the non-rigid parameter is a change amount about a position change of the predetermined feature point due to the motion of the portion including the predetermined feature point.

19. The image processing method according to any one of Claims 11 to 18, wherein the motion model is based on rotation and translation motions of the 3D object, and wherein the third parameter is a translation parameter for specifying a translation amount of the 3D object and a rotation parameter for specifying a rotation amount of the 3D object.

20. The image processing method according to any one of Claims 11 to 19, wherein extended Kalman filtering is applied as said filtering.

21. An image processing program for letting a computer operate to estimate a motion of a predetermined feature point of a 3D object from a motion picture of the 3D object taken by a monocular camera, the image processing program letting the computer execute:

a 3D model initialization step of extracting initial projected coordinates in a model coordinate arithmetic expression for calculation of model coordinates of the predetermined feature point on the basis of a first parameter, a second parameter, and the initial projected coordinates, from an initial frame of the motion picture; and

a motion estimation step of calculating estimates of state variables including a third parameter in a motion arithmetic expression for calculation of coordinates of the predetermined feature point at a time of photography when a processing target frame of the motion picture different from the initial frame was taken, from the model coordinates, the first parameter, and the second parameter, and outputting an output value about the motion of the predetermined feature point on the basis of the second parameter included in the

estimates of the state variables,

wherein the model coordinate arithmetic expression is based on back projection of the monocular camera, the first parameter is a parameter independent of a local motion of a portion including the predetermined feature point, and the second parameter is a parameter dependent on the local motion of the portion including the predetermined feature point,

the image processing program letting the computer operate so that the motion estimation step comprises:

calculating predicted values of the state variables at the time of photography when the processing target frame was taken, based on a state transition model;

applying the initial projected coordinates, and the first parameter and the second parameter included in the predicted values of the state variables, to the model coordinate arithmetic expression to calculate estimates of the model coordinates at the time of photography;

applying the third parameter in the predicted values of the state variables and the estimates of the model coordinates to the motion arithmetic expression to calculate estimates of coordinates of the predetermined feature point at the time of photography;

applying the estimates of the coordinates of the predetermined feature point to an observation function based on an observation model of the monocular camera to calculate estimates of an observation vector of the predetermined feature point;

extracting projected coordinates of the predetermined feature point from the processing target frame, as the observation vector; and

filtering the predicted values of the state variables by use of the extracted observation vector and the estimates of the observation vector to calculate estimates of the state variables at the time of photography.

22. A computer-readable recording medium comprising a record of an image processing program for letting a computer operate to estimate a motion of a predetermined feature point of a 3D object from a motion picture of the 3D object taken by a monocular camera, the image processing program letting the computer execute:

a 3D model initialization step of extracting initial projected coordinates in a model coordinate arithmetic expression for calculation of model coordinates of the predetermined feature point on the basis of a first parameter, a second parameter, and the initial projected coordinates, from an initial frame of the motion picture; and

a motion estimation step of calculating estimates of state variables including a third parameter in a motion arithmetic expression for calculation of coordinates of the predetermined feature point at a time of photography when a processing target frame of the motion picture different from the initial frame was taken, from the model coordinates, the first parameter, and the second parameter, and outputting an output value about the motion of the predetermined feature point on the basis of the second parameter included in the estimates of the state variables,

wherein the model coordinate arithmetic expression is based on back projection of the monocular camera, the first parameter is a parameter independent of a local motion of a portion including the predetermined feature point, and the second parameter is a parameter

dependent on the local motion of the portion including the predetermined feature point,

the image processing program letting the computer operate so that the motion estimation step comprises:

calculating predicted values of the state variables at the time of photography when the processing target frame was taken, based on a state transition model;

applying the initial projected coordinates, and the first parameter and the second parameter included in the predicted values of the state variables, to the model coordinate arithmetic expression to calculate estimates of the model coordinates at the time of photography;

applying the third parameter in the predicted values of the state variables and the estimates of the model coordinates to the motion arithmetic expression to calculate estimates of coordinates of the predetermined feature point at the time of photography;

applying the estimates of the coordinates of the predetermined feature point to an observation function based on an observation model of the monocular camera to calculate estimates of an observation vector of the predetermined feature point;

extracting projected coordinates of the predetermined feature point from the processing target frame, as the observation vector; and

filtering the predicted values of the state variables by use of the extracted observation vector and the estimates of the observation vector to calculate estimates of the state variables at the time of photography.

23. An image processing apparatus for taking a picture of a face with a monocular camera and determining a gaze from the motion

picture thus taken,

wherein a 3D structure of a center of a pupil on the facial picture is defined by a static parameter and a dynamic parameter, and wherein the gaze is determined by estimating the static parameter and the dynamic parameter.

24. The image processing apparatus according to Claim 23, wherein the static parameter is a depth of the pupil in a camera coordinate system.

25. The image processing apparatus according to Claim 23 or Claim 24, wherein the dynamic parameter is a rotation parameter of an eyeball.

26. The image processing apparatus according to Claim 25, wherein the rotation parameter of the eyeball has two degrees of freedom to permit rotations with respect to two coordinate axes in an eyeball coordinate system.

27. An image processing apparatus for taking a picture of a 3D object with a monocular camera and determining a motion of the 3D object from the motion picture thus taken,

wherein a 3D structure of the 3D object on the picture is defined by a rigid parameter and a non-rigid parameter and wherein the motion of the 3D object is determined by estimating the rigid parameter and the non-rigid parameter.

28. The image processing apparatus according to Claim 27, wherein the rigid parameter is a depth of a feature point of the 3D object in a model coordinate system.

29. The image processing apparatus according to Claim 27 or

Claim 28, wherein the non-rigid parameter is a change amount of a feature point of the 3D object in a model coordinate system.

30. An image processing method of taking a picture of a face with a monocular camera and determining a gaze from the motion picture thus taken,

the image processing method comprising:

defining a 3D structure of a center of a pupil on the facial picture by a static parameter and a dynamic parameter; and

determining the gaze by estimating the static parameter and the dynamic parameter.

31. The image processing method according to Claim 30, wherein the static parameter is a depth of the pupil in a camera coordinate system.

32. The image processing method according to Claim 30 or Claim 31, wherein the dynamic parameter is a rotation parameter of an eyeball.

33. The image processing method according to Claim 32, wherein the rotation parameter of the eyeball has two degrees of freedom to permit rotations with respect to two coordinate axes in an eyeball coordinate system.

34. An image processing method of taking a picture of a 3D object with a monocular camera and determining a motion of a 3D object from the motion picture thus taken,

the image processing method comprising:

defining a 3D structure of the 3D object on the picture by a rigid parameter and a non-rigid parameter, and determining the motion of the

3D object by estimating the rigid parameter and the non-rigid parameter.

35. The image processing method according to Claim 34, wherein the rigid parameter is a depth of a feature point of the 3D object in a model coordinate system.

36. The image processing method according to Claim 34 or Claim 35, wherein the non-rigid parameter is a change amount of a feature point of the 3D object in a model coordinate system.

37. An image processing program for letting a computer operate to determine a gaze from a motion picture of a face taken by a monocular camera, the image processing program letting the computer operate to:

define a 3D structure of a center of a pupil on the facial picture by a static parameter and a dynamic parameter, and determine the gaze by estimating the static parameter and the dynamic parameter.

38. An image processing program for letting a computer operate to determine a motion of a 3D object from a motion picture of the 3D object taken by a monocular camera, the image processing program letting the computer operate to:

define a 3D structure of the 3D object on the picture by a rigid parameter and a non-rigid parameter, and determine the motion of the 3D object by estimating the rigid parameter and the non-rigid parameter.

39. A computer-readable recording medium comprising a record of an image processing program for letting a computer operate to determine a gaze from a motion picture of a face taken by a monocular camera, the image processing program being read by the computer to make the computer operate to:

define a 3D structure of a center of a pupil on the facial picture by a static parameter and a dynamic parameter, and determine the gaze by estimating the static parameter and the dynamic parameter.

40. A computer-readable recording medium comprising a record of an image processing program for letting a computer operate to determine a motion of a 3D object from a motion picture of the 3D object taken by a monocular camera, the image processing program being read by the computer to make the computer operate to:

define a 3D structure of the 3D object on the picture by a rigid parameter and a non-rigid parameter, and determine the motion of the 3D object by estimating the rigid parameter and the non-rigid parameter.